

INVERSE PROBLEMS IN ADDITIVE NUMBER THEORY AND IN NON-ABELIAN GROUP THEORY

G. A. FREIMAN, M. HERZOG, P. LONGOBARDI, M. MAJ, AND Y. V. STANCHESCU

1. INTRODUCTION

The aim of this paper is threefold:

- a) Finding new direct and inverse results in the additive number theory concerning Minkowski sums of dilates.
- b) Finding a connection between the above results and some direct and inverse problems in the theory of Baumslag-Solitar (non-abelian) groups.
- c) Solving certain inverse problems in Baumslag-Solitar groups or monoids, assuming appropriate small doubling properties.

We start with our first topic (a), concerning the additive number theory. In this paper \mathbb{Z} denotes the rational integers, \mathbb{N} denotes the **non-negative** elements of \mathbb{Z} and the size of a finite set A will be denoted by $|A|$. Subsets of \mathbb{Z} of the form

$$r * A = \{rx : x \in A\},$$

where r is a positive integer and A is a **finite** subsets of \mathbb{Z} , are called *r-dilates*.

Minkowski sums of dilates are defined as follows:

$$r_1 * A + \dots + r_s * A = \{r_1 x_1 + \dots + r_s x_s : x_i \in A, 1 \leq i \leq s\},$$

These sums have been recently studied in different situations by Nathanson, Bukh, Cilleruelo, Silva, Vinuesa, Hamidoune, Serra and Rué (see [9], [1], [3], [2], [7]). In particular, they examined sums of two dilates of the form

$$A + r * A = \{a + rb \mid a, b \in A\}$$

and solved various *direct* and *inverse* problems concerning their sizes.

For example, it was shown in [3] that

$$|A + 2 * A| \geq 3|A| - 2,$$

Keywords: Inverse additive number theory; Minkowski sums of dilates; Baumslag-Solitar groups and monoids, inverse non-abelian group theory, extended inverse problems.

Mathematics Subject Classification 2010: Primary 11P70; Secondary 11B30, 20F05, 20F99, 11B13, 05E99.

which represents a *direct* result. Moreover, they solved the following *inverse* problem: what is the *structure* of the set A if

$$|A + 2 * A| = 3|A| - 2?$$

Their answer was that in such case A must be an arithmetic progression.

Inverse problems of this type, where the exact bound is assumed, will be called *ordinary inverse problems*. The term *extended inverse problem* will refer to inverse problems in which a small diversion from the exact bound is allowed, still enabling us to reach a definite conclusion concerning the *structure* of A .

As an example of an extended inverse problem, consider the following question: what is the structure of the set A if

$$|A + 2 * A| < 4|A| - 4?$$

Our answer to this question is:

(A). *If $|A + 2 * A| < 4|A| - 4$, then A is a **subset** of an arithmetic progression of size $\leq 2|A| - 3$.* (see Theorem 4, Section 3)

The above mentioned authors and others studied also the sums $A + r * A$ for $r \geq 3$. In this direction we proved the following new (direct) result:

(B). *If $r \geq 3$, then $|A + r * A| \geq 4|A| - 4$.* (see Theorem 6, Section 5)

This very useful result yields a **uniform** bound for all sets A and for $r \geq 3$. In the literature, most bounds of this type are asymptotic.

We continue now with the second topic (b), dealing with a connection, noticed by us, between results concerning sums of dilates and some problems in the theory of Baumslag-Solitar groups.

If S and T are subsets of a group G , their *product* is defined as follows:

$$ST = \{st \mid s \in S, t \in T\}.$$

In particular, $S^2 = \{s_1s_2 \mid s_1, s_2 \in S\}$ and if $b \in G$, then $bS = \{bs \mid s \in S\}$.

For integers m and n , the general Baumslag-Solitar group $BS(m, n)$ is a group with two generators a, b and one defining relation $b^{-1}a^mb = a^n$:

$$BS(m, n) = \langle a, b \mid a^mb = ba^n \rangle.$$

We shall concentrate on

$$G = BS(1, n) = \langle a, b \mid ab = ba^n \rangle.$$

Let S be a finite subset of G of size k_1 contained in the coset $b^r\langle a \rangle$ for some $r \in \mathbb{N}$ and let T be a finite subset of G of size k_2 contained in the coset $b^s\langle a \rangle$ for some $s \in \mathbb{N}$. Then

$$S = \{b^ra^{x_0}, b^ra^{x_1}, \dots, b^ra^{x_{k_1-1}}\},$$

where $A = \{x_0, x_1, \dots, x_{k_1-1}\}$ is a subset of \mathbb{Z} . We introduce now the notation

$$S = \{b^ra^x : x \in A\} = b^ra^A.$$

Thus $|S| = |A|$.

Similarly, $T = b^s a^B$ for some subset $B = \{y_0, y_1, \dots, y_{k_2-1}\}$ of \mathbb{Z} . Since $ab = ba^n$, it follows that $a^{-1}b = ba^{-n}$ and

$$a^x b^t = b^t a^{n^t x} \quad \text{for each } x \in \mathbb{Z} \quad \text{and } t \in \mathbb{N}. \quad (1)$$

In particular,

$$a^x b = b a^{n^x} \quad \text{for each } x \in \mathbb{Z}.$$

Equation (1) implies that

$$(b^r a^x)(b^s a^y) = b^r (a^x b^s) a^y = b^r (b^s a^{n^s x}) a^y = b^{r+s} a^{n^s x + y}$$

for each $x, y \in \mathbb{Z}$ and for each $r, s \in \mathbb{N}$. Therefore the *product set*

$$ST = \{vw \mid v \in S, w \in T\}$$

can be written as

$$\begin{aligned} ST &= \{(b^r a^{x_i})(b^s a^{y_j}) \mid i \in \{0, 1, \dots, k_1 - 1\}, j \in \{0, 1, \dots, k_2 - 1\}\} \\ &= \{b^{r+s} a^{n^s x_i + y_j} \mid i \in \{0, 1, \dots, k_1 - 1\}, j \in \{0, 1, \dots, k_2 - 1\}\} = b^{r+s} a^{n^s * A + B} \end{aligned} \quad (2)$$

and $|ST| = |n^s * A + B|$.

We have proved the following **basic** theorem.

Theorem 1. *Suppose that*

$$S = b^r a^A \subseteq BS(1, n), \quad T = b^s a^B \subseteq BS(1, n)$$

where $r, s \in \mathbb{N}$ and A, B are finite subsets of \mathbb{Z} . Then

$$ST = b^{r+s} a^{n^s * A + B}$$

and

$$|ST| = |n^s * A + B|.$$

In particular,

$$S^2 = b^{2r} a^{n^r * A + A}$$

and

$$|S^2| = |n^r * A + A|.$$

This result will serve us as the major means for investigating $|ST|$, and in particular $|S^2|$, using information about sizes of sums of dilates.

Skipping to our third topic (c), dealing with inverse problems in Baumslag-Solitar groups, it follows from Theorem 1 and from the results mentioned in topic (a), that, using the previous notation, the following statements hold:

(C). *If $S = b a^A \subseteq BS(1, 2)$, then $|S^2| = |2 * A + A|$. Hence $|S^2| \geq 3|S| - 2$ and if $|S^2| = 3|S| - 2$, then A is an arithmetic progression.* (see Theorem 2(a), Section 2)

(D). *If $S = b a^A \subseteq BS(1, 2)$ and $|S^2| < 4|S| - 4$, then A is a **subset** of an arithmetic progression of size $\leq 2|S| - 3$.* (see Theorem 5, Section 4)

(E). If $S = ba^A \subseteq BS(1, r)$ with $r \geq 3$, then $|S^2| \geq 4|S| - 4$. (see Corollary 1, Section 5)

(F). If $S = b^m a^A \subseteq BS(1, 2)$ with $m \geq 2$ an integer, then $|S^2| \geq 4|S| - 4$. (see Corollary 2, Section 5)

For more results concerning S^2 , when $S = ba^A \subseteq BS(1, n)$, see Section 2.

Conditions of the type $|S^2| < 4|S| - 4$ are called *small doubling property*. Our final and main result deals with **arbitrary** finite non-abelian subsets S of the monoid $BS^+(1, 2)$, satisfying the small doubling property $|S^2| < 3.5|S| - 4$. This monoid is defined as follows:

$$BS^+(1, 2) = \{g = b^m a^x \in BS(1, 2) : m, x \in \mathbb{Z}, m \geq 0\}$$

and it is a subset of $BS(1, 2)$, which is closed with respect to multiplication.

We proved the following **general** result concerning subsets of $BS^+(1, 2)$ (see Theorem 7 in Section 6).

(G). If S is a finite non-abelian subset of $BS^+(1, 2)$ satisfying

$$|S^2| < 3.5|S| - 4,$$

then (i) $|S^2| \geq 3|S| - 2$, (ii) $S = ba^A$ for some finite subset A of \mathbb{Z} , which is **contained** in an arithmetic progression of size $< 1.5|S| - 2$ and (iii) $|S^2| = 3|S| - 2$ implies that A is an arithmetic progression of length $|S|$.

Our paper is a pilot study in the following more general direction. Let G be an infinite non-abelian group of certain type and let S denote a finite **non-abelian** subset (i.e. $\langle S \rangle$ is non-abelian) of G of order k (k -subset in short). It is natural to ask the following questions:

Q.1. Find $m_G(k)$, the minimal possible value of $|S^2|$ for non-abelian k -subsets S of G .

Q.2. What can we say about the detailed structure of *extremal k -subsets* of G , i.e. finite non-abelian subsets S of G of size k , satisfying

$$|S^2| = m_G(k)?$$

Q.3. More generally, what can we say about the detailed structure of non-abelian k -subsets S of G , satisfying some *small doubling property*, say,

$$m_G(k) \leq |S^2| < c_0 k + d_0,$$

where c_0 is a small constant greater than 1 and d_0 is some small constant.

As mentioned above, we tried to answer these questions in the case of $G = BS(1, n)$ and in particular for $G = BS(1, 2)$. We hope that our work will lead to similar studies for other classes of non-abelian groups.

This paper is a contribution to the current programme of extending the Freiman-type theory, concerning the structure of subsets of \mathbb{Z} with the small doubling property, to such subsets of non-abelian groups (see, for example, [4], [6] and [14]).

In this paper we use the following notation. We write $[m, n] = \{x \in \mathbb{Z} \mid m \leq x \leq n\}$. The *algebraic sum* of two finite subsets A and B of \mathbb{Z} will be denoted by

$$A + B = \{a + b \mid a \in A, b \in B\}.$$

In particular, if $b \in \mathbb{Z}$, then $A + b = \{a + b \mid a \in A\}$. The sum $2A = A + A$ is called the *sumset* of A . Throughout this paper we shall use the well known inequality

$$|A + B| \geq |A| + |B| - 1.$$

Let $A = \{a_0 < a_1 < \dots < a_{k-1}\}$ be a finite increasing set of k integers. By the *length* $\ell(A)$ of A we mean the difference

$$\ell(A) = \max(A) - \min(A) = a_{k-1} - a_0$$

between its maximal and minimal elements and

$$h_A = \ell(A) + 1 - |A|$$

denotes the number of *holes* in A , that is $h_A = |[a_0, a_{k-1}] \setminus A|$. Finally, if $k \geq 2$, then we denote

$$d(A) = g.c.d.(a_1 - a_0, a_2 - a_0, \dots, a_{k-1} - a_0).$$

We shall use several times the following result of Lev-Smelianski and Stanchescu:

Theorem LSS. *Let A and B be finite subsets of \mathbb{N} such that $0 \in A \cap B$. Define*

$$\delta_{A,B} = \begin{cases} 1, & \text{if } \ell(A) = \ell(B), \\ 0, & \text{if } \ell(A) \neq \ell(B). \end{cases}$$

Then the following statements hold:

(i) *If $\ell(A) = \max(\ell(A), \ell(B)) \geq |A| + |B| - 1 - \delta_{A,B}$ and $d(A) = 1$, then*

$$|A + B| \geq |A| + 2|B| - 2 - \delta_{A,B}.$$

(ii) *If $\max(\ell(A), \ell(B)) \leq |A| + |B| - 2 - \delta_{A,B}$, then*

$$|A + B| \geq (|A| + |B| - 1) + \max(h_A, h_B) = \max(\ell(A) + |B|, \ell(B) + |A|).$$

Proof. Assertion (i) is Theorem 2(ii) from [8]. Assertion (ii) is Theorem 4 from [11]. \square

2. EXTREMAL SETS CONTAINED IN ONE COSET OF $BS(1, n)$

In this section we consider finite subsets S of

$$G = BS(1, n) = \langle a, b \mid ab = ba^n \rangle$$

which are contained in the coset $b\langle a \rangle$ of $\langle a \rangle$ in G . In other words, if $|S| = k$, then

$$S = b\{a^{x_0}, a^{x_1}, \dots, a^{x_{k-1}}\} = ba^A,$$

where $A = \{x_0, x_1, \dots, x_{k-1}\} \subseteq \mathbb{Z}$.

In view of Theorem 1, questions Q.1 and Q.2 concerning such S belong to the *additive number theory*: find a tight lower bound for the size of the Minkowski sum $n * A + A$ and describe the structure of extremal sets A .

For $n = 2$ and $n = 3$, the answer to questions Q.1 and Q.2 are known. Using Theorems 1.1 and 1.2 in [3] and Theorem 1, we get the following group-theoretical results:

Theorem 2. *Let $A \subseteq \mathbb{Z}$ be a finite set of integers. Then the following statements hold.*

- (a) *If $S = ba^A \subseteq BS(1, 2)$, then $|S^2| \geq 3|S| - 2$. Moreover, equality holds if and only if A is an arithmetic progression.*
- (b) *If $S = ba^A \subseteq BS(1, 3)$, then*

$$|S^2| \geq 4|S| - 4.$$

Moreover, equality holds if and only if either one of the following holds:

$$A = \{0, 1, 3\}, \quad A = \{0, 1, 4\}, \quad A = 3 * \{0, \dots, n\} \cup (3 * \{0, \dots, n\} + 1)$$

or A is an affine transform of one of these sets.

Proof. (a) It follows from Theorems 1.1 in [3] that $|A + 2 * A| \geq 3|A| - 2$ and $|A + 2 * A| = 3|A| - 2$ if and only if A is an arithmetic progression. Since $|S^2| = |A + 2 * A|$ by Theorem 1, we get (a).

(b) It follows from Theorem 1.2 in [3] that $|A + 3 * A| \geq 4|A| - 4$ and $|A + 3 * A| = 4|A| - 4$ if and only if either one of the following holds:

$$A = \{0, 1, 3\}, \quad A = \{0, 1, 4\}, \quad A = 3 * \{0, \dots, n\} \cup (3 * \{0, \dots, n\} + 1)$$

or A is an affine transform of one of these sets. Since $|S^2| = |A + 3 * A|$ by Theorem 1, we get (b). \square

For $n \geq 4$, Theorem 1 and known results concerning sums of dilates yield the following partial results.

Theorem 3. *Let $A \subseteq \mathbb{Z}$ be a finite set of integers and let $S = ba^A$ be a subset of $BS(1, n)$. Then:*

- (a) *If $S \subseteq BS(1, 4)$ and $|S| \geq 5$, then $|S^2| \geq 5|S| - 6$.*
- (b) *If $S \subseteq BS(1, n)$, then $|S^2| \geq (n + 1)|S| - o(|S|)$.*
- (c) *If p is an odd prime number, $S \subseteq BS(1, p)$ and $|S| \geq 3(p - 1)^2(p - 1)!$, then*

$$|S^2| \geq (p + 1)|S| - \lceil k(k + 2)/4 \rceil.$$

*Moreover, equality holds if and only if $A = p * \{0, \dots, m\} + \{0, \dots, \frac{p-1}{2}\}$ for some m .*

Proof. Using Theorem 1, we get $|S^2| = |n * A + A|$.

Inequality (a) follows from $|S^2| = |4 * A + A|$ and Theorem 3 in [10].

Inequality (b) follows from $|S^2| = |n * A + A|$ and Theorem 1.2 in [1].

Assertion (c) follows from $|S^2| = |p * A + A|$ and Corollary 1.3 in [2].

□

3. AN EXTENDED INVERSE RESULT FOR $|A + 2 * A|$.

In this section we extend Theorem 1.1 in [3], which states that $|A + 2 * A| \geq 3|A| - 2$ for any finite subset A of \mathbb{Z} and $|A + 2 * A| = 3|A| - 2$ implies that A is an arithmetic progression. In Theorem 4 below, we prove the following *extended inverse result* in the additive number theory: if A is a finite subset of \mathbb{Z} of size $|A| \geq 3$ satisfying $|A + 2 * A| < 4|A| - 4$, then A is **contained** in an arithmetic progression of size $2|A| - 3$ at most. This result will be used in the next section.

Theorem 4. *Let $A = \{a_0 < a_1 < a_2 < \dots < a_{k-1}\} \subset \mathbb{Z}$ be a finite set of integers of size $k = |A| \geq 1$. Then the following statements hold.*

- (a) *If $1 \leq k \leq 2$, then $|A + 2 * A| = 3k - 2$ and A is an arithmetic progression of size k .*
- (b) *If $k \geq 3$, assume that*

$$|A + 2 * A| = (3k - 2) + h < 4k - 4. \quad (3)$$

Then

$$h \geq 0, \quad |A + 2 * A| \geq 3k - 2$$

*and the set A is a **subset** of an arithmetic progression*

$$P = \{a_0, a_0 + d, a_0 + 2d, \dots, a_0 + (l - 1)d\}$$

of size $|P|$ bounded by

$$|P| \leq k + h = |A + 2 * A| - 2k + 2 \leq 2k - 3. \quad (4)$$

- (c) *If $k \geq 1$ and $|A + 2 * A| = 3k - 2$, then A is an arithmetic progression*

$$A = \{a_0, a_0 + d, a_0 + 2d, \dots, a_0 + (k - 1)d\}.$$

Proof. (a) If $k = 1$, then $|A + 2 * A| = 1 = 3k - 2$ and A is an arithmetic progression of size k . If $k = 2$ and $A = \{a < b\}$, then

$$A + 2 * A = \{3a, a + 2b, b + 2a, 3b\}.$$

Since $a \neq b$, it follows that $|A + 2 * A| = 4 = 3k - 2$ and A is an arithmetic progression of size k . The proof of (a) is complete.

- (b) We assume now that $k \geq 3$ and (3) holds. Suppose, first, that A is *normal*, i.e.

$$\min(A) = a_0 = 0 \text{ and } d = d(A) = \gcd(A) = 1. \quad (5)$$

Thus $\ell(A) = a_{k-1}$.

We split the set A into a disjoint union

$$A = A_0 \cup A_1,$$

where $A_0 \subseteq 2\mathbb{Z}$ and $A_1 \subseteq 2\mathbb{Z} + 1$. Since $0 = a_0 \in A_0$ and $d(A) = 1$, it follows that $A_0 \neq \emptyset$ and $A_1 \neq \emptyset$. Therefore

$$m = |A_0| \geq 1, \quad n = |A_1| \geq 1 \text{ and } k = m + n.$$

We denote

$$A_0 = \{0 = 2x_0 < 2x_1 < \dots < 2x_{m-1}\}, \quad A_0^* = \frac{1}{2}A_0 = \{0 < x_1 < \dots < x_{m-1}\},$$

$$A_1 = \{2y_0 + 1 < 2y_1 + 1 < \dots < 2y_{n-1} + 1\},$$

and

$$A_1^* = \frac{1}{2}(A_1 - 1) - y_0 = \{0 < y_1 - y_0 < y_2 - y_0 < \dots < y_{n-1} - y_0\}.$$

Thus

$$\ell(A_0^*) = x_{m-1} < a_{k-1} = \ell(A) \text{ and also } \ell(A_1^*) = y_{n-1} - y_0 < a_{k-1} = \ell(A).$$

The set $A + 2 * A$ is the union of two disjoint subsets $A_0 + 2 * A \subseteq 2\mathbb{Z}$ and $A_1 + 2 * A \subseteq 2\mathbb{Z} + 1$ and therefore

$$|A + 2 * A| = |A_0 + 2 * A| + |A_1 + 2 * A| = |A_0^* + A| + |A_1^* + A|. \quad (6)$$

We continue our proof with two claims.

Claim 1:

$$\ell(A) \leq k + \max(m, n) - 2 \leq 2k - 3. \quad (7)$$

For the proof of Claim 1 we shall use Theorem LSS (i). Since $\ell(A) > \ell(A_0^*), \ell(A_1^*)$, we have $\delta_{A, A_0^*} = \delta_{A, A_1^*} = 0$.

Suppose, first, that $m \leq n$. If the claim is false, then

$$\ell(A) \geq k + n - 1 = |A| + |A_1^*| - 1 \geq k + m - 1 = |A| + |A_0^*| - 1$$

and since $d(A) = 1$, Theorem LSS (i) yields the following inequalities:

$$|A_0^* + A| \geq k + 2|A_0^*| - 2 = k + 2m - 2 \text{ and } |A_1^* + A| \geq k + 2|A_1^*| - 2 = k + 2n - 2. \quad (8)$$

Using (6) and (8), we get that $|A + 2 * A| \geq 4k - 4$, which contradicts our hypothesis (3).

Similarly, if $n \leq m$ and

$$\ell(A) \geq k + m - 1 \geq k + n - 1,$$

then $d(A) = 1$ and Theorem LSS (i) imply again the inequalities (8), which together with (6) yield $|A + 2 * A| \geq 4k - 4$, a contradiction.

Hence $\ell(A) \leq k + \max(m, n) - 2$. Since $k = m + n$ and $m, n \geq 1$, it follows that $\max(m, n) \leq k - 1$ and hence $\ell(A) \leq k + \max(m, n) - 2 \leq 2k - 3$. The proof of Claim 1 is complete.

Next we state and prove Claim 2.

Claim 2:

$$|A + 2 * A| \geq (3k - 2) + h_A. \quad (9)$$

Recall that $h_A = \ell(A) + 1 - |A|$. For the proof of Claim 2 we shall use Claim 1 and Theorem LSS(ii). We distinguish between two cases.

Case 1: Suppose that $m \leq n$ and hence, by (7), $\ell(A) \leq k + n - 2$.

Thus it follows by Theorem LSS(ii) that

$$|A_1^* + A| \geq (n + k - 1) + h_A$$

and therefore

$$\begin{aligned} |A + 2 * A| &= |A_0^* + A| + |A_1^* + A| \\ &\geq (|A_0^*| + |A| - 1) + |A_1^* + A| \geq (m + k - 1) + (n + k - 1) + h_A \\ &= (3k - 2) + h_A. \end{aligned}$$

Case 2: Suppose that $n < m$ and hence, by (7), $\ell(A) \leq k + m - 2$.

Thus it follows by Theorem LSS(ii) that

$$|A_0^* + A| \geq (m + k - 1) + h_A$$

and therefore

$$|A + 2 * A| = |A_0^* + A| + |A_1^* + A| \geq (m + k - 1) + h_A + (n + k - 1) = (3k - 2) + h_A.$$

In both cases we obtain that h_A , the total number of holes in the normal set A , satisfies

$$0 \leq h_A \leq |A + 2 * A| - (3k - 2) = h \leq k - 3.$$

Hence

$$h \geq h_A \geq 0 \quad \text{and} \quad |A + 2 * A| \geq (3k - 2).$$

Moreover, the set A is contained in the arithmetic progression

$$P = \{a_0, a_0 + 1, a_0 + 2, \dots, a_{k-1}\} = \{0, 1, 2, \dots, a_{k-1}\}$$

of size

$$a_{k-1} + 1 = k + h_A \leq k + h \leq 2k - 3. \quad (10)$$

It follows that Theorem 4 (b) holds for *normal* sets A satisfying (5) and (3).

Let now A be an *arbitrary* finite set of $k = |A| \geq 3$ integers satisfying the inequality (3). We define

$$B = \frac{1}{d(A)}(A - a_0) = \left\{ \frac{1}{d(A)}(x - a_0) : x \in A \right\}.$$

Note that $|B| = |A| = k$, $\min(B) = 0$, $d(B) = 1$ and

$$|B + 2 * B| = |A + 2 * A| = (3k - 2) + h < 4k - 4.$$

Therefore B is a normal set satisfying inequality (3) of Theorem 4 and as shown above

$$0 \leq h_B \leq |B + 2 * B| - (3k - 2) = |A + 2 * A| - (3k - 2) = h \leq k - 3.$$

Hence also in the general case we get

$$h \geq 0 \quad \text{and} \quad |A + 2 * A| \geq (3k - 2).$$

Moreover, it follows from (10) applied to B that B is contained in the arithmetic progression

$$P = \{0, 1, 2, \dots, b_{k-1}\}$$

with

$$b_{k-1} = \max(B) \leq k + h - 1 \leq 2k - 4.$$

Thus $A = d(A)B + a_0$ is contained in an arithmetic progression

$$\{a_0, a_0 + d, a_0 + 2d, \dots, a_0 + (k + h - 1)d\}$$

of size $k + h \leq 2k - 3$, where d denotes $d(A)$. The proof of (b) is complete.

(c) If $1 \leq k \leq 2$, then our claim follows from (a). So suppose that $k \geq 3$. Then $h = 0$ and by (4) in (b), A is a subset of an arithmetic progression of size k at most. But A is a set of size k , so A is equal to the arithmetic progression. The proof of (c), and hence also of Theorem 4, is now complete. \square

4. AN EXTENDED INVERSE RESULT FOR SUBSETS OF $b\langle a \rangle$ IN $BS(1, 2)$.

In this section we shall apply Theorem 4 in order to obtain an *extended inverse result* in group theory.

Recall that $BS(1, 2) = \langle a, b \mid ab = ba^2 \rangle$. In Theorem 2 we obtained the following inverse group-theoretical result:

If $A \subseteq \mathbb{Z}$ is a finite set of integers and $S = ba^A \subset BS(1, 2)$, then

$$|S^2| \geq 3|S| - 2.$$

Moreover, equality holds if and only if A is an arithmetic progression.

Theorem 4, together with Theorem 1, allow us to solve the corresponding extended inverse group-theoretical problem.

Theorem 5. *Let $A \subseteq \mathbb{Z}$ be a finite set of integers of size $k = |A| \geq 1$. If $S = ba^A$ is a finite subset of the group $BS(1, 2)$, then $|S| = k$ and*

$$|S^2| \geq 3k - 2. \tag{11}$$

Moreover, if $k \geq 3$ and

$$|S^2| = (3k - 2) + h < 4|S| - 4, \tag{12}$$

then $h \geq 0$ and S is a subset of a geometric progression

$$S \subseteq \{ba^u, ba^{u+d}, ba^{u+2d}, \dots, ba^{u+(k+h-1)d}\}$$

of size $k + h \leq 2k - 3$, where $u = \min(A)$ and $d = d(A)$.

Furthermore, if either $1 \leq k \leq 2$ or $k \geq 3$ and $h = 0$, then S is the geometric progression

$$S = \{ba^u, ba^{u+d}, ba^{u+2d}, \dots, ba^{u+(k-1)d}\}.$$

Proof. Clearly $|S| = |A| = k$ and by Theorem 1, $|S^2| = |2 * A + A|$. Hence it follows by Theorem 4 that $|S^2| \geq 3k - 2$, proving (11).

If $k \geq 3$, then (12) implies, again by Theorem 1, that

$$|A + 2 * A| = (3k - 2) + h < 4k - 4.$$

Hence it follows by Theorem 4, that $h \geq 0$ and A is a subset of an arithmetic progression

$$P = \{u, u + d, u + 2d, \dots, u + (k + h - 1)d\}$$

of size $k + h \leq 2k - 3$, where $u = \min(A)$ and $d = d(A)$. Hence

$$S \subseteq \{ba^u, ba^{u+d}, ba^{u+2d}, \dots, ba^{u+(k+h-1)d}\}.$$

Finally, if either $1 \leq k \leq 2$ or $k \geq 3$ and $h = 0$, then, by Theorem 4, A is an arithmetic progression and hence S is the required geometric progression. \square

5. A NEW LOWER BOUND FOR $|A + r * A|$ AND APPLICATIONS.

In this section we obtain a new tight lower bound for $|A + r * A|$, provided that $r \geq 3$.

Theorem 6. *Let $A = \{a_0 < a_1 < a_2 < \dots < a_{k-1}\} \subset \mathbb{Z}$ be a finite set of integers of size $|A| = k \geq 1$. Then for every integer $r \geq 3$ we have*

$$|A + r * A| \geq \max(4k - 4, 1) \geq 3k - 2. \quad (13)$$

Remark. If $r = 3$, then Theorem 6 follows from Theorem 1.2 in [3]. If $r \geq 4$, then the results of [1] and [2] are asymptotically stronger than (13), but we need a lower bound valid for *every* k . Our proof is independent of [3].

Proof. If $k = 1$, then $|A + r * A| = 1 = \max(4k - 4, 1) = 3k - 2$ and the theorem holds.

If $k = 2$, then $A = \{a < b\}$ and $r > 1$ implies that $a + rb \neq b + ra$. Hence

$$|A + r * A| = |\{a, b\} + \{ra, rb\}| = |\{(r+1)a, b + ra, a + rb, (r+1)b\}| = 4 = 4k - 4 = 3k - 2,$$

so the theorem holds also for $k = 2$. Therefore we shall assume, from now on, that $k \geq 3$. Thus, since $k > 1$, we need only to prove that

$$|A + r * A| \geq 4k - 4.$$

We assume first that A is *normal*, i.e.

$$\min(A) = a_0 = 0 \quad \text{and} \quad d = d(A) = \gcd(A) = 1. \quad (14)$$

We split the set A into a disjoint union of s non-empty subsets, each of which being contained in a **distinct** residue class modulo r :

$$A = A_1 \cup A_2 \cup \dots \cup A_s,$$

where

$$A_i \subseteq x_i + r\mathbb{Z}, \quad |A_i| \geq 1 \quad \text{and} \quad x_i = \min A_i.$$

Note that $k \geq 3$, $d(A) = 1$ and $\min(A) = a_0 = 0$, so $s \geq 2$.

We clearly have

$$|A + r * A| = \sum_{i=1}^s |A_i + r * A| \geq \sum_{i=1}^s (|A_i| + |A| - 1) = |A| + s(|A| - 1).$$

If $s \geq 3$, then we get $|A + r * A| \geq 4|A| - 3$ and Theorem 6 follows.

Hence we may assume that $s = 2$ and $A = A_1 \cup A_2$, where A_1 and A_2 are non-empty subsets of A contained in disjoint residue classes modulo r . Let

$$k_1 = |A_1| \quad \text{and} \quad k_2 = |A_2|.$$

Then $k = k_1 + k_2$ and we may assume, without loss of generality, that

$$k_1 \geq k_2.$$

Hence $2k_1 \geq k$ and $k_1 \geq 2$.

Recall that if S is a finite subset of \mathbb{Z} , then $\ell(S)$, the length of S , is defined by $\ell(S) = \max(S) - \min(S)$. For $i = 1, 2$ we define

$$A_i^* = \frac{1}{r}(A_i - \min(A_i)) = \left\{ \frac{1}{r}(x - x_i) : x \in A_i \right\}.$$

Clearly $|A_i^*| = |A_i|$ and we have

$$|A_i + r * A| = |A_i^* + A|.$$

Thus

$$|A + r * A| = |A_1 + r * A| + |A_2 + r * A| = |A_1^* + A| + |A_2^* + A|.$$

Note also that

$$\ell(A_i) \geq r(k_i - 1) \quad \text{and} \quad \ell(A_i^*) = \frac{1}{r}\ell(A_i),$$

so

$$k_i - 1 \leq \ell(A_i^*) = \frac{1}{r}\ell(A_i) \leq \ell(A_i) \leq \ell(A).$$

Moreover, $\ell(A_i) > \ell(A_i^*)$ if and only if $k_i > 1$, so $\ell(A_1) > \ell(A_1^*)$ since $k_1 \geq 2$.

Clearly we must have either $k_1 = k - 1 > k_2 = 1$ or $k_1 \geq k_2 > 1$. We shall examine these two cases separately.

Case 1: Suppose that $k_1 = k - 1 > k_2 = 1$. We have $k = k_1 + 1$ and $\ell(A) \geq \ell(A_1) > \ell(A_1^*)$. Moreover,

$$\ell(A) \geq \ell(A_1) \geq r(k_1 - 1) \geq 3k_1 - 3.$$

We distinguish now between two complementary subcases.

(i) Suppose that $\ell(A) \geq k + k_1 - 1 = 2k_1$. Then, since $d(A) = 1$, Theorem LSS(i) implies that

$$|A + A_1^*| \geq k + 2k_1 - 2.$$

(ii) Suppose that $\ell(A) \leq k + k_1 - 2 = 2k_1 - 1$. Then, since $k_1 \geq 2$, Theorem LSS(ii) implies that

$$|A + A_1^*| \geq \ell(A) + |A_1| \geq 3k_1 - 3 + k_1 = 4k_1 - 3 \geq 3k_1 - 1 = k + 2k_1 - 2.$$

Thus in both cases we have

$$|A + r * A| = |A_1^* + A| + |A_2^* + A| \geq (k + 2k_1 - 2) + k = 4k - 4,$$

as required

Case 2: Suppose that $k_1 \geq k_2 > 1$. Then

$$\ell(A) > \ell(A_1^*), \quad \ell(A) > \ell(A_2^*)$$

and

$$\ell(A) \geq \ell(A_i) \geq r(k_i - 1) \geq 3k_i - 3$$

for $i = 1, 2$. We distinguish now between three complementary subcases.

(i) Suppose that $\ell(A) \geq k + k_1 - 1$. Then also $\ell(A) \geq k + k_2 - 1$ and since $d(A) = 1$, Theorem LSS(i) implies that

$$|A + A_1^*| \geq k + 2k_1 - 2, \quad |A + A_2^*| \geq k + 2k_2 - 2.$$

Hence

$$|A + r * A| = |A_1^* + A| + |A_2^* + A| \geq (k + 2k_1 - 2) + (k + 2k_2 - 2) = 4k_1 + 4k_2 - 4 = 4k - 4,$$

as required.

(ii) Suppose that $k + k_2 - 1 \leq \ell(A) \leq k + k_1 - 2$. Then

$$k_1 \geq k_2 + 1$$

and since $d(A) = 1$, Theorem LSS(i) and (ii) imply that

$$|A + A_1^*| \geq \ell(A) + |A_1^*| \geq 3k_1 - 3 + k_1 = 4k_1 - 3 \text{ and } |A + A_2^*| \geq k + 2k_2 - 2.$$

Hence

$$|A + r * A| = |A_1^* + A| + |A_2^* + A| \geq 5k_1 + 3k_2 - 5 \geq 4k_1 + 4k_2 - 4 = 4k - 4,$$

as required.

(iii) Suppose that $\ell(A) \leq k + k_2 - 2$. Then $3k_1 - 3 \leq \ell(A) \leq k_1 + 2k_2 - 2$, yielding $2k_1 \leq 2k_2 + 1$. Since $k_1 \geq k_2$, it follows that

$$k_1 = k_2 \geq 2$$

and

$$3k_i - 3 \leq r(k_i - 1) \leq \ell(A_i) \leq \ell(A) \leq k + k_2 - 2 = 3k_1 - 2 = 3k_2 - 2.$$

We claim that $\ell(A) = 3k_1 - 2$. Indeed, if $\ell(A) = 3k_1 - 3$, then $\ell(A_1) = \ell(A_2) = \ell(A) = a_{k-1}$. But $a_{k-1} \notin A_i$ for some i and hence $\ell(A_i) < a_{k-1}$, a contradiction. This proves our claim.

Recall that $\ell(A) > \ell(A_1^*)$ and $\ell(A) > \ell(A_2^*)$. Since $\ell(A) = 3k_1 - 2 = k + (k_i - 2) = |A| + |A_i^*| - 2$ for $i = 1, 2$, it follows, by Theorem LSS(ii), that

$$|A + r * A| = |A_1^* + A| + |A_2^* + A| \geq \ell(A) + k_1 + \ell(A) + k_2 = 2(3k_1 - 2) + k = 4k - 4,$$

as required. Our proof in Case 2 is complete.

So Theorem 6 holds for *normal* sets A . Let A be now an *arbitrary* finite set of $k = |A| \geq 3$ integers. We define

$$B = \frac{1}{d(A)}(A - a_0) = \left\{ \frac{1}{d(A)}(x - a_0) : x \in A \right\}.$$

Note that $|B| = |A| = k$, $\min(B) = 0$, $d(B) = 1$ and $|A + r * A| = |B + r * B|$. For the normal set B we have proved that $|B + r * B| \geq 4|B| - 4$. It follows that

$$|A + r * A| = |B + r * B| \geq 4|B| - 4 = 4|A| - 4,$$

as required. The proof of Theorem 6 is complete. \square

Theorem 6 yields the following two applications. Here is the first one.

Corollary 1. *Let $S \subseteq BS(1, r)$ be a finite set of size $k = |S| \geq 1$ and suppose that $r \geq 3$ and*

$$S = ba^A,$$

where $A \subseteq \mathbb{Z}$ is a finite set of integers.

Then

$$|S^2| = |A + r * A| \geq \max(4k - 4, 1) \geq 3k - 2. \quad (15)$$

Proof. By Theorem 1, $|S^2| = |A + r * A|$ and hence, by Theorem 6, $|S^2| \geq \max(4k - 4, 1) \geq 3k - 2$, as required. \square

Our next application will be used several times in the proof of the main Theorem 7 in Section 6.

Corollary 2. *Let $S \subseteq BS(1, 2)$ be a finite set of size $k = |S| \geq 1$ and suppose that*

$$S = b^m a^A,$$

where $m \geq 2$ is an integer and $A \subseteq \mathbb{Z}$ is a finite set of integers.

Then

$$S^2 = b^{2m} a^{A+2^m * A}. \quad (16)$$

and

$$|S^2| = |A + 2^m * A| \geq \max(4k - 4, 1) \geq 3k - 2. \quad (17)$$

Proof. By Theorem 1, $|S^2| = |A + 2^m * A|$. Since $2^m > 3$, it follows by Corollary 1 that $|S^2| \geq \max(4k - 4, 1) \geq 3k - 2$, as required. \square

6. AN EXTENDED INVERSE RESULT FOR ALL SUBSETS OF $BS^+(1, 2)$.

In Section 5 we proved an extended inverse result for finite subsets of $BS(1, 2)$ which are contained in the coset $ba^{\mathbb{Z}}$. In this section we solve, using a more detailed analysis, a more general problem concerning **all** finite non-abelian subsets S of the corresponding monoid

$$BS^+(1, 2) = \{g = b^m a^x \in BS(1, 2) \mid x, m \in \mathbb{Z}, m \geq 0\}, \quad (18)$$

which satisfy the more restrictive small doubling property:

$$|S^2| < 3.5|S| - 4.$$

We proved the following theorem.

Theorem 7. *If S be a finite non-abelian subset of $BS^+(1, 2)$ of size $|S| = k$, then*

$$|S^2| \geq 3k - 2. \quad (19)$$

Moreover, if

$$|S^2| = (3k - 2) + h < 3.5k - 4, \quad (20)$$

then there exists a finite set of integers $A \subseteq \mathbb{Z}$ such that

- (a) $S = ba^A$
- (b) *The set A is contained in an arithmetic progression of size*

$$k + h < 1.5k - 2.$$

Throughout this section we shall use the following **notation**. $BS^+(1, 2)$ is the monoid defined by (18). Every element $g \in BS^+(1, 2)$ can be represented in a *unique way* as a product

$$g = b^m a^x,$$

where $m \in \mathbb{N}$ and $x \in \mathbb{Z}$. It follows that for every two distinct natural numbers $m \neq n$, we have

$$b^m a^{\mathbb{Z}} \cap b^n a^{\mathbb{Z}} = \emptyset. \quad (21)$$

If

$$S \subseteq BS^+(1, 2)$$

is a finite subset of $BS^+(1, 2)$ of size $k = |S|$, we define a set of natural numbers

$$M_S = M(S) \subseteq \mathbb{N}$$

by the following condition: $m \in M_S$ if and only if there is an integer x such that $b^m a^x \in S$. The set S defines M_S in a unique way and we will denote it by

$$M_S = \{m_0 < m_1 < \dots < m_t\},$$

where $t \geq 0$ and $m_0 \geq 0$. For every $0 \leq i \leq t$, we define

$$S_i = S \cap b^{m_i} a^{\mathbb{Z}}, \quad k_i = |S_i|. \quad (22)$$

Every set S_i is non-empty, lies in only one coset of the cyclic subgroup $\langle a \rangle = a^{\mathbb{Z}}$ and there is a finite set of integers $A_i \subseteq \mathbb{Z}$ such that

$$S_i = b^{m_i} a^{A_i} \subseteq b^{m_i} a^{\mathbb{Z}}.$$

The set S can be written as a *disjoint union* of $t + 1$ sets

$$S = S_0 \cup S_1 \cup \dots \cup S_t, \quad (23)$$

satisfying

$$k_i = |S_i| = |A_i| \geq 1.$$

Example 1. Theorem 7 is *optimal* in view of the following example:

$$S = a^{A_0} \cup \{b\} \subset BS^+(1, 2),$$

where

$$A_0 = \{0, 1, 2, \dots, k-2\} \text{ and } k \text{ is even.}$$

The set S is clearly non-abelian and

$$S^2 = a^{A_0} a^{A_0} \cup b a^{A_0} \cup a^{A_0} b \cup \{b^2\}.$$

Using $a^{A_0} b = b a^{2 \cdot A_0}$, we get

$$S^2 = a^{A_0 + A_0} \cup (b a^{A_0} \cup b a^{2 \cdot A_0}) \cup \{b^2\} = a^{A_0 + A_0} \cup b a^{A_0 \cup 2 \cdot A_0} \cup \{b^2\}.$$

Since

$$a^{A_0 + A_0} \subseteq a^{\mathbb{Z}}, \quad b a^{A_0 \cup 2 \cdot A_0} \subseteq b a^{\mathbb{Z}}, \quad \{b^2\} \subseteq b^2 a^{\mathbb{Z}},$$

it follows by (21) that the three components of S^2 are disjoint in pairs and hence

$$|S^2| = |A_0 + A_0| + |A_0 \cup 2 \cdot A_0| + 1 = (2k - 3) + (1.5k - 2) + 1 = 3.5k - 4. \quad (24)$$

This example shows that if $|S^2| \geq 3.5k - 4$, then we have to take into account sets that are not included in only one coset of the cyclic subgroup $\langle a \rangle$ generated by a .

The proof of Theorem 7 will follow from Lemmas 1-7 below.

Lemma 1. *Let $S \subseteq BS^+(1, 2)$ be a finite set of size $k = |S|$. Suppose that $t \geq 1$ and there is $0 \leq j \leq t$ such that $k_j = |S_j| \geq 2$. Then S generates a non-abelian group.*

Proof. If $j = 0$ and $m_0 = 0$, then $k_0 = |S_0| = |A_0| \geq 2$ implies that $S_0 \neq \{1\}$ and $A_0 \neq \{0\}$. Since $t \geq 1$, it follows that there are three integers m, x, z such that $m \geq 1, x \neq 0, a^x \in S_0$ and $b^m a^z \in S_1$. In this case

$$a^x(b^m a^z) = b^m a^{z+2^m x} \neq (b^m a^z)a^x = b^m a^{z+x}$$

and therefore S generates a non-abelian group.

It remains to examine the following two cases:

- (i) $j \geq 1$.
- (ii) $j = 0$ and $m_0 \geq 1$.

If $j \geq 1$, then $m_j \geq 1$ and $k_j = |S_j| = |b^{m_j} a^{A_j}| \geq 2$ implies that $|A_j| \geq 2$. On the other hand, if $j = 0$ and $m_0 \geq 1$, then $k_0 = |S_0| = |b^{m_0} a^{A_0}| \geq 2$ implies that $|A_0| \geq 2$. In both cases, let $m = m_j$. Then $m \geq 1$ and there are two integers $x \neq y$ such that $\{b^m a^x, b^m a^y\} \subseteq S_j$. We conclude that

$$(b^m a^x)(b^m a^y) = b^{2m} a^{y+2^m x} \neq (b^m a^y)(b^m a^x) = b^{2m} a^{x+2^m y},$$

since $x \neq y$ and $m \geq 1$. The proof of Lemma 1 is complete. \square

We shall examine now the case $t = 1$, i.e. we shall study sets S lying in exactly two cosets. Note that inequality (25) in the following Lemma 2 is tight, in view of Example 1.

Lemma 2. *Let $S \subseteq BS^+(1, 2)$ be a finite set of size $k = |S| \geq 2$. Suppose that $S = U \cup V$ with $U = b^m a^M \neq \emptyset$ and $V = b^n a^N \neq \emptyset$, where $0 \leq m < n$ are two integers and $M, N \subseteq \mathbb{Z}$. Then*

$$|S^2| \geq 3.5k - 4. \quad (25)$$

Proof. Clearly $k = |M| + |N|$ and

$$S^2 = U^2 \cup (UV \cup VU) \cup V^2. \quad (26)$$

Using Theorem 1 we get

$$U^2 = b^{2m} a^{M+2^m * M}, \quad V^2 = (b^n a^N)(b^n a^N) = b^{2n} a^{N+2^n * N}, \quad (27)$$

$$UV = (b^m a^M)(b^n a^N) = b^{m+n} a^{N+2^m * M}, \quad VU = (b^n a^N)(b^m a^M) = b^{m+n} a^{M+2^n * N}. \quad (28)$$

Since the sets $b^{2m} a^{\mathbb{Z}}$, $b^{m+n} a^{\mathbb{Z}}$ and $b^{2n} a^{\mathbb{Z}}$ are disjoint in pairs, it follows that

$$|S^2| = |U^2| + |(UV \cup VU)| + |V^2|. \quad (29)$$

We shall examine now two complementary cases.

Case 1: $1 \leq m < n$.

We shall estimate $|U^2|$ and $|V^2|$ using either Theorem 5 or Corollary 2. We have

$$|U^2| = |M + 2^m * M| \geq 3|M| - 2, \quad |V^2| = |N + 2^n * N| \geq 3|N| - 2.$$

Using (29) and $|UV| = |N + 2^n * M| \geq |M| + |N| - 1$ we conclude that

$$|S^2| \geq |U^2| + |UV| + |V^2| \geq 3|M| - 2 + (|M| + |N| - 1) + 3|N| - 2 = 4k - 5 \geq 3.5k - 4,$$

as required.

Case 2: $0 = m < n$.

In this case S is a disjoint union of two non-empty sets:

$$S = U \cup V, \text{ where } U = a^M, V = b^n a^N \text{ and } n \geq 1.$$

We have

$$U^2 = a^{M+M}, V^2 = b^{2n} a^{N+2^n * N}, \quad (30)$$

$$UV = b^n a^{N+2^n * M}, VU = b^n a^{M+N}. \quad (31)$$

Therefore it follows, either by Theorem 5 or by Corollary 2, that

$$|U^2| = |M + M|, \quad |V^2| = |N + 2^n * N| \geq 3|N| - 2. \quad (32)$$

We also clearly have

$$\begin{aligned} |UV \cup VU| &= |(N + 2^n * M) \cup (M + N)| = |(M \cup 2^n * M) + N| \\ &\geq |(M \cup 2^n * M)| + |N| - 1 \geq |M| + |N| - 1. \end{aligned} \quad (33)$$

Suppose that $|M| = 1$. Then it follows from (29), (33) and (32) that

$$|S^2| \geq 1 + (1 + |N| - 1) + (3|N| - 2) = 4|N| - 1 \geq 3.5(1 + |N|) - 4 = 3.5|S| - 4,$$

as required. So we may assume that $|M| \geq 2$.

We shall complete the proof by dealing separately with two complementary sub-cases. Denote

$$\ell = \ell(M) = \max(M) - \min(M), \quad d = d(M) = \gcd\{x - \min(M) : x \in M\}$$

and define

$$M^* = \frac{1}{d}(M - \min(M)), \quad \ell^* = \ell(M^*) = \max(M^*) = \frac{\ell}{d}.$$

Case 2.1. Assume that $\ell(M^*) \geq 2|M^*| - 2$.

As shown above, we may assume that $|M| \geq 2$. Suppose that $|M| = 2$. Then $M = \{a_0 < a_1\}$, which implies that $d(M) = a_1 - a_0$ and $M^* = \{0, 1\}$. Thus $\ell(M^*) = 1$ and by our assumptions $1 = \ell(M^*) \geq 2|M^*| - 2 = 2$, a contradiction. Hence we may assume that $|M| \geq 3$, which implies that $k = |M| + |N| \geq 3 + 1 = 4$.

Note that $d(M^*) = 1$. By using Theorem LSS(i) for equal summands we get

$$|U^2| = |M + M| = |M^* + M^*| \geq 3|M^*| - 3 = 3|M| - 3. \quad (34)$$

Using (29), (34), (32) and (33), we may conclude that

$$|S^2| \geq |U^2| + |UV \cup VU| + |V^2| \geq (3|M| - 3) + (|M| + |N| - 1) + (3|N| - 2) = 4k - 6.$$

Since $k \geq 4$, it follows that $|S^2| \geq 3.5k - 4$, as required.

Case 2.2. Assume that $\ell(M^*) \leq 2|M^*| - 3$.

In this case, we use Theorem LSS(ii) for equal summands. Let $h_{M^*} = \ell^* + 1 - |M^*|$ be the number of *holes* in M^* . We get

$$|M + M| = |M^* + M^*| \geq 2|M^*| - 1 + h_{M^*} = |M^*| + \ell^* = |M| + \ell^*. \quad (35)$$

We shall now estimate the size of $M \cap 2^n * M$. Note that all the common elements of $2^n * M$ and M lie in the interval $[\min(M), \max(M)]$ of length ℓ and the set $2^n * M$ is included in an arithmetic progression of difference $2^n d \geq 2d$. Therefore

$$|M \cap (2^n * M)| \leq \frac{\ell}{2d} + 1 = \frac{\ell^*}{2} + 1 \quad (36)$$

and

$$|M \cup (2^n * M)| = |M| + |2^n * M| - |M \cap 2^n * M| \geq 2|M| - \frac{\ell^*}{2} - 1. \quad (37)$$

Using (29), (32), (33), (35) and (37) we conclude that

$$\begin{aligned} |S^2| &\geq |U^2| + |UV \cup VU| + |V^2| \geq \\ &\geq |M + M| + (|M \cup 2^n * M| + |N| - 1) + (|N + 2^n * N|) \\ &\geq (|M| + \ell^*) + (2|M| - \frac{\ell^*}{2} - 1 + |N| - 1) + (3|N| - 2) \\ &= 3|M| + 4|N| - 4 + \frac{\ell^*}{2} \\ &\geq 3|M| + 4|N| - 4 + \frac{|M^*| - 1}{2} = 3.5|M| + 4|N| - 4.5 \geq 3.5k - 4, \end{aligned} \quad (38)$$

as required. \square

In Lemmas 3,4,5,6 we shall obtain tight lower bounds for the cardinality of $|S^2|$, assuming that $k_i = |S_i| \geq 2$ for at most one i , $0 \leq i \leq t$.

Lemma 3. *Let $S \subseteq BS^+(1, 2)$ be a finite set of size $k = |S|$. Suppose that*

$$S = S_0 \cup S_1 \cup \dots \cup S_t, \quad (39)$$

where $t \geq 2$. If $k_0 = |S_0| \geq 2$ and $k_i = |S_i| = 1$ for every $1 \leq i \leq t$, then

$$|S^2| \geq 4k - 5 > 3.5k - 4. \quad (40)$$

Example 2. Inequality (40) is tight.

If

$$S = \{1, a\} \cup \{b, b^2, \dots, b^t\},$$

then $k = t + 2$ and

$$S^2 = \{1, a, a^2\} \cup \{b, b^2, \dots, b^t\} \cup \{ab, ab^2, \dots, ab^t\} \cup \{ba, b^2a, \dots, b^ta\} \cup \{b^2, b^3, \dots, b^{2t}\}.$$

Note that equality (1) implies that

$$\{ab, ab^2, \dots, ab^t\} = \{ba^2, b^2a^4, \dots, b^ta^{2^t}\}$$

and thus

$$S^2 = \{1, a, a^2\} \cup \bigcup_{j=1}^t b^j \{1, a, a^{2^j}\} \cup \{b^{t+1}, b^{t+2}, \dots, b^{2t}\}.$$

Using (21), we get $|S^2| = 3(t+1) + t = 4t + 3 = 4k - 5$. □

We continue now with the proof of Lemma 3.

Proof. Clearly $k = k_0 + t \geq 2 + 2 = 4$. Let

$$A_0 = \{y_1 < \dots < y_{k_0}\} \subset \mathbb{Z}$$

be a finite set of k_0 integers that defines the set

$$S_0 = b^{m_0} a^{A_0} = \{b^{m_0} a^{y_1}, \dots, b^{m_0} a^{y_{k_0}}\}$$

with $k_0 \geq 2$, and let

$$S_i = \{b^{m_i} a^{x_i}\}$$

for every $1 \leq i \leq t$. Recall our assumption that $0 \leq m_0 < m_1 < \dots < m_t$.

Note that for every $1 \leq i \leq t$ we have $m_i > 0$,

$$S_0 S_i = b^{m_0+m_i} \{a^{x_i+2^{m_i}y_1}, \dots, a^{x_i+2^{m_i}y_{k_0}}\}, \quad |S_0 S_i| = k_0$$

and

$$S_i S_0 = b^{m_i+m_0} \{a^{y_1+2^{m_0}x_i}, \dots, a^{y_{k_0}+2^{m_0}x_i}\}, \quad |S_i S_0| = k_0.$$

We claim that

$$|S_0 S_i \cup S_i S_0| \geq k_0 + 1. \tag{41}$$

Indeed, if $S_0 S_i = S_i S_0$, then

$$\{x_i + 2^{m_i}y_1 < \dots < x_i + 2^{m_i}y_{k_0}\} = \{y_1 + 2^{m_0}x_i < \dots < y_{k_0} + 2^{m_0}x_i\}$$

and thus

$$(2^{m_0} - 1)x_i = (2^{m_i} - 1)y_1 = \dots = (2^{m_i} - 1)y_{k_0},$$

which contradicts $\{y_1 < \dots < y_{k_0}\}$, in view of $m_i \geq 1$ and $k_0 \geq 2$.

Note that

$$S_0 S_i \cup S_i S_0 \subseteq b^{m_0+m_i} a^{\mathbb{Z}}, \quad S_i S_t \subseteq b^{m_i+m_t} a^{\mathbb{Z}},$$

for every $0 \leq i \leq t$. Moreover, $S_0 S_0 = b^{2m_0} a^{A_0+2^{m_0}*A_0}$, so $|S_0 S_0| = |A_0 + 2^{m_0} * A_0| \geq 2|A_0| - 1 = 2k_0 - 1$. It follows by (21) that the sets

$$S_0 S_0, S_0 S_1 \cup S_1 S_0, \dots, S_0 S_t \cup S_t S_0, S_1 S_t, \dots, S_t S_t$$

are disjoint and included in S^2 . Using $t \geq 2$, $k_0 \geq 2$, (41) and $k \geq 4$, we conclude that

$$\begin{aligned} |S^2| &\geq (|S_0S_0| + |S_0S_1 \cup S_1S_0| + \dots + |S_0S_t \cup S_tS_0|) + (|S_1S_t| + \dots + |S_tS_t|) \\ &\geq (2k_0 - 1) + (k_0 + 1) + \dots + (k_0 + 1) + (1 + \dots + 1) = (2k_0 - 1) + t(k_0 + 1) + t \\ &= 4k_0 + (t - 2)k_0 + 2t - 1 \geq 4k_0 + 2(t - 2) + 2t - 1 = 4k_0 + 4t - 5 = 4k - 5 \\ &> 3.5k - 4, \end{aligned} \tag{42}$$

as required. \square

Lemma 4. *Let $S \subseteq BS^+(1, 2)$ be a finite set of size $k = |S|$. Suppose that*

$$S = S_0 \cup S_1 \cup \dots \cup S_t, \tag{43}$$

where $t \geq 2$. If $k_t = |S_t| \geq 2$ and $k_i = |S_i| = 1$ for every $0 \leq i \leq t - 1$, then

$$|S^2| \geq 4k - 5 > 3.5k - 4. \tag{44}$$

Example 3. Inequality (44) is tight.

If

$$S = \{1, b, b^2, \dots, b^{t-1}\} \cup \{b^t, b^t a\},$$

then $k = t + 2$ and

$$S^2 = \{1, b, b^2, \dots, b^{2t-1}\} \cup \{1, b, \dots, b^{t-1}\} b^t a \cup b^t a \{1, b, \dots, b^{t-1}\} \cup \{b^{2t}, b^{2t} a, b^t a b^t, b^t a b^t a\}.$$

Note that equality (1) implies that

$$b^t a \{1, b, \dots, b^{t-1}\} = \{b^t a, b^{t+1} a^2, \dots, b^{2t-1} a^{2^{t-1}}\}$$

and

$$\{b^{2t}, b^{2t} a, b^t a b^t, b^t a b^t a\} = \{b^{2t}, b^{2t} a, b^{2t} a^{2^t}, b^{2t} a^{2^{t+1}}\}.$$

Thus

$$S^2 = \{1, b, b^2, \dots, b^{t-1}\} \cup b^t \{1, a\} \cup \bigcup_{j=1}^{t-1} b^{t+j} \{1, a, a^{2^j}\} \cup b^{2t} \{1, a, a^{2^t}, a^{2^{t+1}}\}$$

and by (21), $|S^2| = t + 2 + 3(t - 1) + 4 = 4t + 3 = 4k - 5$. \square

We continue now with the proof of Lemma 4.

Proof. Clearly $k = k_t + t \geq 2 + 2 = 4$. Let

$$A_t = \{y_1 < \dots < y_{k_t}\} \subseteq \mathbb{Z}$$

be a finite set of $k_t \geq 2$ integers, which defines the set

$$S_t = b^{m_t} a^{A_t} = \{b^{m_t} a^{y_1}, \dots, b^{m_t} a^{y_{k_t}}\}$$

and let

$$S_i = \{b^{m_i} a^{x_i}\}$$

for every $0 \leq i \leq t-1$.

Note that for every $0 \leq i \leq t-1$ we have

$$S_t S_i = b^{m_t+m_i} \{a^{x_i+2^{m_i}y_1}, \dots, a^{x_i+2^{m_i}y_{k_t}}\}, \quad |S_t S_i| = k_t$$

and

$$S_i S_t = b^{m_i+m_t} \{a^{y_1+2^{m_t}x_i}, \dots, a^{y_{k_t}+2^{m_t}x_i}\}, \quad |S_i S_t| = k_t.$$

It follows, like in Lemma 3, that

$$|S_t S_i \cup S_i S_t| \geq k_t + 1 \quad (45)$$

for $1 \leq i \leq t-1$.

Note that $|S_t S_t| \geq 3k_t - 2$, in view of Corollary 2. Using $|S_0 S_t| = k_t$, $k_t \geq 2$, (45) and $k \geq 4$, we conclude, like in Lemma 3, that

$$\begin{aligned} |S^2| &\geq (|S_0 S_0| + \dots + |S_0 S_t|) + (|S_1 S_t \cup S_t S_1| + \dots + |S_{t-1} S_t \cup S_t S_{t-1}| + |S_t S_t|) \\ &\geq (1 + \dots + 1 + k_t) + ((t-1)(k_t + 1) + (3k_t - 2)) = 4k_t + (t-1)k_t + 2t - 3 \\ &\geq 4k_t + 4t - 5 = 4k - 5 > 3.5k - 4, \end{aligned} \quad (46)$$

as required. \square

Lemma 5. *Let $S \subseteq BS^+(1, 2)$ be a finite non-abelian set of size $k = |S| \geq 2$. Suppose that*

$$S = S_0 \cup S_1 \cup \dots \cup S_t \quad (47)$$

where $|S_i| = 1$ for all i and

$$S_i = \{s_i\}, \quad s_i = b^{m_i} a^{x_i}.$$

Denote $T = S \setminus \{s_0\}$. If the subgroup $\langle T \rangle$ is abelian, then

$$|S^2| \geq 4k - 4. \quad (48)$$

Proof. Recall that $M_S = \{m_0 < m_1 < \dots < m_t\}$, where $m_0 \geq 0$. We notice first that $T \neq \emptyset$ since $k \geq 2$. Moreover, we claim that the sets T^2 , $s_0 T \cup T s_0$ and $\{s_0^2\}$ are disjoint. Indeed, we have:

- (i) $s_0^2 \notin T^2$, since $s_0^2 = (b^{m_0} a^{x_0})^2 = b^{2m_0} a^{x_0+2^{m_0}x_0}$ and $T^2 \subseteq \{b^m a^x : m \geq 2m_1\}$.
- (ii) $s_0^2 \notin (s_0 T \cup T s_0)$, because $s_0 \notin T$.
- (iii) $s_0 \notin \langle T \rangle$, because $\langle T \rangle$ is abelian and $\langle S \rangle$ is non-abelian. This implies that $s_0 T \cup T s_0$ does not intersect the set T^2 .

Notice also that $|T| = t$ and if $s_i, s_j \in T$, then $s_i s_j = b^{m_i+m_j} a^{2^{m_j}x_i+x_j}$, which implies that $|T^2| \geq |M_S \setminus \{m_0\} + M_S \setminus \{m_0\}| \geq 2|M_S \setminus \{m_0\}| - 1 = 2t - 1$.

In order to complete the proof of Lemma 5, it suffices to show that the sets s_0T and Ts_0 are disjoint. Indeed, if that is the case, then

$$\begin{aligned} |S^2| &\geq |T^2| + |s_0T \cup Ts_0| + |\{s_0^2\}| \\ &= |T^2| + |s_0T| + |Ts_0| + |\{s_0^2\}| \\ &\geq (2t - 1) + t + t + 1 = 4t = 4|S| - 4, \end{aligned}$$

as required.

So suppose, by way of contradiction, that

$$s_0T \cap Ts_0 \neq \emptyset. \quad (49)$$

Note that

$$s_0T = \{s_0s_1, \dots, s_0s_t\} = \{b^{m_0+m_1}a^{x_1+2^{m_1}x_0}, \dots, b^{m_0+m_t}a^{x_t+2^{m_t}x_0}\},$$

and

$$Ts_0 = \{s_1s_0, \dots, s_ts_0\} = \{b^{m_0+m_1}a^{x_0+2^{m_0}x_1}, \dots, b^{m_0+m_t}a^{x_0+2^{m_0}x_t}\}.$$

Therefore (49) implies that there is $1 \leq i \leq t$ such that

$$s_0s_i = b^{m_0+m_i}a^{x_i+2^{m_i}x_0} = s_is_0 = b^{m_0+m_i}a^{x_0+2^{m_0}x_i}$$

and thus

$$(2^{m_i} - 1)x_0 = (2^{m_0} - 1)x_i. \quad (50)$$

Choose an arbitrary $1 \leq j \leq t$. Since $\langle T \rangle$ is abelian, it follows that

$$s_js_i = b^{m_j+m_i}a^{x_i+2^{m_i}x_j} = s_is_j = b^{m_j+m_i}a^{x_j+2^{m_j}x_i},$$

yielding

$$(2^{m_i} - 1)x_j = (2^{m_j} - 1)x_i.$$

Hence

$$x_i = \frac{2^{m_i} - 1}{2^{m_j} - 1}x_j$$

and from (50) we get

$$(2^{m_i} - 1)x_0 = (2^{m_0} - 1)\frac{2^{m_i} - 1}{2^{m_j} - 1}x_j.$$

That means that $(2^{m_j} - 1)x_0 = (2^{m_0} - 1)x_j$ and thus

$$s_0s_j = b^{m_0+m_j}a^{x_j+2^{m_j}x_0} = b^{m_0+m_j}a^{x_0+2^{m_0}x_j} = s_js_0.$$

It follows that s_0 commutes with every element of T , which contradicts our assumption that $\langle T \rangle$ is abelian and $\langle S \rangle$ is non-abelian. The proof of Lemma 5 is complete. \square

Lemma 6. *Let $S \subseteq BS^+(1, 2)$ be a finite set of cardinality $k = |S| \geq 2$. Suppose that S is a disjoint union*

$$S = V_1 \cup V_2 \cup \dots \cup V_t, \quad (51)$$

of t subsets

$$V_i = \{s_i\}, s_i = b^{m_i} a^{x_i},$$

of size $|V_i| = 1$. If S is a non-abelian set and $1 \leq m_1 < m_2 < \dots < m_t$, then

$$|S^2| \geq 4k - 4. \quad (52)$$

Example 4. Inequality (52) is tight.

If $S = \{b, b^2, \dots, b^{t-1}\} \cup \{b^t a\}$, then $k = t$ and S^2 is the union of four disjoint sets:

$$\{b^2, b^3, \dots, b^{2t-2}\},$$

$$\{b, b^2, \dots, b^{t-1}\} b^t a = \{b^{t+1} a, b^{t+2} a, \dots, b^{2t-1} a\},$$

$$b^t a \{b, b^2, \dots, b^{t-1}\} = \{b^{t+1} a^2, b^{t+2} a^4, \dots, b^{2t-1} a^{2^{t-1}}\}$$

and $\{b^t a b^t a\} = \{b^{2t} a^{2^t+1}\}$. Therefore

$$|S^2| = (2t - 3) + (t - 1) + (t - 1) + 1 = 4k - 4.$$

□

We continue now with the proof of Lemma 6.

Proof. If a set S satisfies all the assumptions of Lemma 6, then we say that S is an *elementary set*.

Clearly $t = k \geq 2$ and we proceed by induction on t . If $t = 2$, then $S = \{s_1, s_2\}$ and since $s_1 s_2 \neq s_2 s_1$ and $s_1^2 \neq s_2^2$, it follows that $|S^2| = 4 = 4|S| - 4$, as required.

For the inductive step, let $t \geq 3$ be an integer, and assume that Lemma 6 holds for each elementary set $T \subseteq BS^+(1, 2)$ of size $2 \leq |T| \leq t - 1$. Denote

$$S' = S \setminus \{s_1\}.$$

In view of Lemma 5, we may assume that $\langle S' \rangle$ is non-abelian.

We shall continue by examining two complementary cases.

Case 1: $s_1 s_2 = s_2 s_1$.

Choose $n \geq 2$ maximal such that the set $S^* := \{s_1, s_2, \dots, s_n\}$ is abelian. Note that $n < t$, because S is a non-abelian set, and $s_{n+1} \notin \langle S^* \rangle$. Moreover, $s_1 s_{n+1} \notin S'^2$, since otherwise $s_1 s_{n+1} = s_u s_v$ for some $2 \leq u, v \leq t$ and hence $b^{m_1+m_{n+1}} = b^{m_u+m_v}$, implying that $m_1 < m_u, m_v < m_{n+1}$, whence $1 < u, v < n + 1$ and $s_{n+1} \in \langle S^* \rangle$, a contradiction. Similarly $s_{n+1} s_1 \notin S'^2$.

We claim that it suffices to show that s_{n+1} does not commute with s_1 .

Indeed, if $s_1 s_{n+1} \neq s_{n+1} s_1$, then (52) follows from

$$\{s_1^2, s_1 s_2, s_1 s_{n+1}, s_{n+1} s_1\} \subseteq S^2 \setminus S'^2$$

and from the induction hypothesis for S' :

$$|S^2| \geq |S'^2| + |\{s_1^2, s_1s_2, s_1s_{n+1}, s_{n+1}s_1\}| \geq (4|S'| - 4) + 4 = 4|S| - 4.$$

We shall complete the proof by showing that if

$$s_1s_{n+1} = s_{n+1}s_1, \tag{53}$$

then

$$s_js_{n+1} = s_{n+1}s_j,$$

for every $1 \leq j \leq n$, which contradicts the maximality of n .

Our argument is similar to that used in the proof of Lemma 5. Denote $m = m_{n+1}$, $x = x_{n+1}$ and

$$s_{n+1} = b^m a^x.$$

We first note that (53) implies that

$$\begin{aligned} s_1s_{n+1} &= (b^{m_1}a^{x_1})(b^m a^x) = b^{m_1+m}a^{x+2^m x_1} \\ &= s_{n+1}s_1 = (b^m a^x)(b^{m_1}a^{x_1}) = b^{m+m_1}a^{x_1+2^{m_1}x} \end{aligned}$$

and thus

$$(2^{m_1} - 1)x = (2^m - 1)x_1. \tag{54}$$

Choose an arbitrary $1 \leq j \leq n$. Using

$$s_1s_j = s_js_1$$

we get, like in the proof of Lemma 5, that

$$x_1 = \frac{2^{m_1} - 1}{2^{m_j} - 1}x_j.$$

It follows by (54) that

$$(2^{m_1} - 1)x = (2^m - 1)\frac{2^{m_1} - 1}{2^{m_j} - 1}x_j$$

and since $m_1 \geq 1$, we may conclude that

$$(2^{m_j} - 1)x = (2^m - 1)x_j.$$

Thus $s_js_{n+1} = s_{n+1}s_j$, a contradiction.

Case 2: $s_1s_2 \neq s_2s_1$.

We *claim* that

$$\text{either } s_1s_3 \neq s_2^2 \text{ or } s_3s_1 \neq s_2^2. \tag{55}$$

Indeed, if

$$s_1s_3 = s_2^2 \quad \text{and} \quad s_3s_1 = s_2^2$$

then

$$(b^{m_1}a^{x_1})(b^{m_3}a^{x_3}) = (b^{m_2}a^{x_2})^2 = (b^{m_3}a^{x_3})(b^{m_1}a^{x_1})$$

and thus

$$b^{m_1+m_3}a^{x_3+2^{m_3}x_1} = b^{2m_2}a^{x_2+2^{m_2}x_2} = b^{m_1+m_3}a^{x_1+2^{m_1}x_3}.$$

It follows that $m_1 + m_3 = 2m_2$ and

$$x_3 = x_2(2^{m_2} + 1) - 2^{m_3}x_1, \quad 2^{m_1}x_3 = x_2(2^{m_2} + 1) - x_1.$$

Thus

$$2^{m_1}x_3 = (2^{m_2} + 1)x_2 - x_1 = 2^{m_1}(2^{m_2} + 1)x_2 - 2^{m_1+m_3}x_1,$$

implying that

$$(2^{m_1} - 1)(2^{m_2} + 1)x_2 = (2^{m_1+m_3} - 1)x_1 = (2^{2m_2} - 1)x_1.$$

Hence

$$(2^{m_1} - 1)x_2 = (2^{m_2} - 1)x_1,$$

and

$$s_1s_2 = b^{m_1+m_2}a^{x_2+2^{m_2}x_1} = b^{m_2+m_1}a^{x_1+2^{m_1}x_2} = s_2s_1,$$

a contradiction. The proof of our claim is complete.

Thus

$$|\{s_1s_3, s_3s_1\} \setminus \{s_2^2\}| \geq 1.$$

Since

$$\{s_1^2, s_1s_2, s_2s_1\} \subseteq S^2 \setminus S'^2$$

and

$$\{s_1s_3, s_3s_1\} \setminus \{s_2^2\} \subseteq S^2 \setminus S'^2,$$

it follows by the induction hypothesis for S' , that

$$|S^2| \geq |S'^2| + |\{s_1^2, s_1s_2, s_2s_1, s_1s_3, s_3s_1\} \setminus S'^2| \geq (4|S'| - 4) + 4 = 4|S| - 4.$$

The proof of Lemma 6 is complete. □

The following lemma is the main step in the proof of Theorem 7. We use an inductive argument analogous to that used for the proof of Lemma 2.2 in [12] (see also Lemma 3 in [13]).

Lemma 7. *Let $S \subseteq BS^+(1, 2)$ be a finite set of size $k = |S| \geq 2$. Suppose that*

$$S = S_0 \cup S_1 \cup \dots \cup S_t, \tag{56}$$

where $t \geq 1$. If S is a non-abelian set, then

$$|S^2| \geq 3.5k - 4. \tag{57}$$

Proof. We use induction on $t \geq 1$. Observe that when $t = 1$, Lemma 7 follows from Lemma 2.

For the inductive step, let $t \geq 2$ be an integer, and assume that Lemma 7 holds for any non-abelian finite set $T \subseteq BS^+(1, 2)$ which lies in u distinct cosets of $\langle a \rangle = a^{\mathbb{Z}}$, where $2 \leq u < t$.

Denote

$$S^* = S \setminus S_t, \quad k^* = |S^*| = k - k_t.$$

If S^* generates a non-abelian group, then our inductive hypothesis implies that

$$|(S^*)^2| \geq 3.5k^* - 4,$$

and it suffices to show that

$$|S_t^2 \cup S_t S_{t-1} \cup S_{t-1} S_t| \geq 3.5k_t, \quad (58)$$

since inequality (57) then follows from

$$|S^2| \geq |(S^*)^2| + |S_t^2 \cup S_t S_{t-1} \cup S_{t-1} S_t| \geq (3.5k^* - 4) + 3.5k_t = 3.5k - 4. \quad (59)$$

The proof of (58) will be provided by examining four complementary cases.

Case 1: Assume that either $k_t \geq 2$ or $k_{t-1} \geq 2$.

Recall that $0 \leq m_0 < m_1 < \dots < m_t$ and hence $t \geq 2$ implies that $m_t \geq 2$. Thus, using either Theorem 5 or Corollary 2, we get

$$|S_t^2| \geq \max\{4k_t - 4, 1\} \geq 3k_t - 2.$$

We shall examine now four subcases.

i. If $k_t + 2k_{t-1} \geq 6$, then (58) is true in view of:

$$|S_t^2 \cup S_t S_{t-1} \cup S_{t-1} S_t| \geq |S_t^2| + |S_t S_{t-1}| \geq (3k_t - 2) + (k_t + k_{t-1} - 1) = 4k_t + k_{t-1} - 3 \geq 3.5k_t.$$

If there is $0 \leq j \leq t-1$ such that $k_j = |S_j| \geq 2$, then S^* generates a non-abelian group (in view of Lemma 1) and we may apply the induction hypothesis. Thus, Lemma 7 follows from (58) and (59).

If $k_j = 1$ for all $0 \leq j \leq t-1$, then $k_t \geq 6 - 2k_{t-1} \geq 4$, in view of (i). In this case, Lemma 7 follows from Lemma 4.

So we may assume that Case (i) does not hold and in particular

$$k_t + 2k_{t-1} \leq 5.$$

Hence one of the following cases must hold: (ii) $k_t = 3, k_{t-1} = 1$, (iii) $k_t = 2, k_{t-1} = 1$ or (iv) $k_t = 1, k_{t-1} = 2$.

ii. If $k_t = 3$ and $k_{t-1} = 1$, then Corollary 2 implies $|S_t^2| \geq 4k_t - 4$ and therefore inequality (58) follows from:

$$|S_t^2 \cup S_t S_{t-1} \cup S_{t-1} S_t| \geq |S_t^2| + |S_t S_{t-1}| \geq (4k_t - 4) + (k_t + k_{t-1} - 1) = 5k_t - 4 > 3.5k_t.$$

If there is $0 \leq j \leq t-1$ such that $k_j = |S_j| \geq 2$, then S^* generates a non-abelian group (in view of Lemma 1) and we may apply the induction hypothesis. Thus, Lemma 7 follows from (58) and (59).

If $k_j = 1$ for all $0 \leq j \leq t-1$, then Lemma 7 follows from Lemma 4, in view of $k_t = 3$.

iii. If $k_t = 2$ and $k_{t-1} = 1$, then we can write

$$S_{t-1} = \{b^u a^x\}, \quad S_t = \{b^v a^y, b^v a^z\},$$

where $1 \leq u = m_{t-1} < v = m_t$ and $y < z$ are integers. Using the identity $a^x b^m = b^m a^{2^m x}$, we get

$$S_{t-1} S_t = b^{u+v} \{a^{2^v x+y}, a^{2^v x+z}\} \quad \text{and} \quad S_t S_{t-1} = b^{u+v} \{a^{2^u y+x}, a^{2^u z+x}\}.$$

Note that $S_{t-1} S_t \neq S_t S_{t-1}$. Indeed, if $S_{t-1} S_t = S_t S_{t-1}$, then $2^v x + y = 2^u y + x$ and $2^v x + z = 2^u z + x$. Thus $(2^v - 1)y = (2^v - 1)x = (2^u - 1)z$, which contradicts $y < z$, in view of $u \geq 1$. Therefore either Theorem 5 or Corollary 2 implies that

$$|S_t^2 \cup S_t S_{t-1} \cup S_{t-1} S_t| = |S_t^2| + |S_t S_{t-1} \cup S_{t-1} S_t| \geq (3k_t - 2) + 3 = 4 + 3 = 3.5k_t.$$

If there is $0 \leq j \leq t-1$ such that $k_j = |S_j| \geq 2$, then S^* generates a non-abelian group (in view of Lemma 1) and we may apply the induction hypothesis. Thus, Lemma 7 follows from (58) and (59).

If $k_j = 1$ for all $0 \leq j \leq t-1$, then Lemma 7 follows from Lemma 4, in view of $k_t = 2$.

iv. If $k_t = 1$ and $k_{t-1} = 2$, then we can write

$$S_{t-1} = \{b^u a^y, b^u a^z\}, \quad S_t = \{b^v a^x\},$$

where $1 \leq u = m_{t-1} < v = m_t$, x, y, z are integers and $y < z$. Using the identity $a^x b^m = b^m a^{2^m x}$, we get

$$S_{t-1} S_t = b^{u+v} \{a^{2^v y+x}, a^{2^v z+x}\} \quad \text{and} \quad S_t S_{t-1} = b^{u+v} \{a^{2^u x+y}, a^{2^u x+z}\}.$$

Note that $S_{t-1} S_t \neq S_t S_{t-1}$. Indeed, if $S_{t-1} S_t = S_t S_{t-1}$, then $2^v y + x = 2^u x + y$ and $2^v z + x = 2^u x + z$. Thus $(2^v - 1)y = (2^u - 1)x = (2^v - 1)z$, which contradicts $y < z$, in view of $v \geq 1$. Therefore,

$$|S_t^2 \cup S_t S_{t-1} \cup S_{t-1} S_t| = |S_t^2| + |S_t S_{t-1} \cup S_{t-1} S_t| \geq 1 + 3 > 3.5k_t.$$

Since $k_{t-1} = 2$, Lemma 1 implies that S^* generates a non-abelian group and we may apply the induction hypothesis. Thus, Lemma 7 follows from (58) and (59).

The proof in **Case 1** is complete.

Case 2: Assume that $k_t = k_{t-1} = \dots = k_1 = 1$ and $k_0 \geq 2$.

In this case, Lemma 7 follows from Lemma 3.

Case 3: Assume that $k_t = k_{t-1} = 1$ and there is $1 \leq j \leq t-2$ such that $k_j \geq 2$ and $k_i = 1$ for every $i \in \{j+1, \dots, t\}$.

Let

$$S_j = b^{m_j} a^{A_j} = \{b^{m_j} a^{y_1}, \dots, b^{m_j} a^{y_{k_j}}\}$$

and let $S_i = \{b^{m_i}a^{x_i}\}$ for every $i \in \{j+1, \dots, t\}$. Clearly $|S_j S_i| = |S_i S_j| = k_j$ and using the reasoning in the proof of Lemma 3, we get

$$|S_j S_i \cup S_i S_j| \geq k_j + 1 \quad (60)$$

for every $i \in \{j+1, \dots, t\}$.

Note that $k_j \geq 2$, so by Lemma 1 the set

$$S_j^* = S_0 \cup S_1 \cup \dots \cup S_j$$

is non-abelian. By applying the inductive hypothesis to S_j^* and in view of (60), we obtain

$$|S^2| \geq |S_j^* S_j^*| + \sum_{u=j+1}^t |S_j S_u \cup S_u S_j| + \sum_{u=j+1}^t |S_u S_t| \quad (61)$$

$$\begin{aligned} &\geq (3.5|S_j^*| - 4) + (k_j + 1)(t - j) + (t - j) \\ &= (3.5|S_j^*| - 4) + (k_j + 2)(t - j) \geq (3.5|S_j^*| - 4) + 4(t - j) \\ &= (3.5|S_j^*| - 4) + 4(k - |S_j^*|) = 4k - 4 - 0.5|S_j^*| > 3.5k - 4, \end{aligned} \quad (62)$$

as required.

Case 4: Assume that $k_i = 1$ for every $0 \leq i \leq t$.

If the set $S' = S \setminus \{s_0\}$ is abelian, then Lemma 5 implies that

$$|S^2| \geq 4|S| - 4 > 3.5k - 4,$$

as required. Therefore, we may assume that $S' = S_1 \cup S_2 \cup \dots \cup S_t$ is non-abelian. Since $t \geq 2$, it follows that $k = t + 1 \geq 3$ and Lemma 6 implies that

$$|S'^2| \geq 4|S'| - 4. \quad (63)$$

Moreover, $\{s_0^2, s_0 s_1\} \subseteq S^2 \setminus S'^2$.

We distinguish now between two complementary cases.

(a) If $k = |S| \geq 4$, then

$$|S^2| \geq |S'^2| + 2 \geq 4(k - 1) - 4 + 2 = 4k - 6 \geq 3.5k - 4,$$

as required.

(b) If $k = 3$, then $S = \{s_0, s_1, s_2\}$, $S' = \{s_1, s_2\}$, $s_1 s_2 \neq s_2 s_1$, $|S'^2| = 4$ and

$$S^2 = \{s_0^2, s_0 s_1, s_1 s_0, s_0 s_2, s_2 s_0\} \cup S'^2.$$

We *claim* that

$$\text{either } s_0 s_2 \neq s_1^2 \text{ or } s_2 s_0 \neq s_1^2. \quad (64)$$

Indeed, if

$$s_0 s_2 = s_1^2 \quad \text{and} \quad s_2 s_0 = s_1^2$$

then

$$(b^{m_0}a^{x_0})(b^{m_2}a^{x_2}) = (b^{m_1}a^{x_1})^2 = (b^{m_2}a^{x_2})(b^{m_0}a^{x_0})$$

and thus

$$b^{m_0+m_2}a^{x_2+2^{m_2}x_0} = b^{2m_1}a^{x_1+2^{m_1}x_1} = b^{m_2+m_0}a^{x_0+2^{m_0}x_2}.$$

It follows that $m_0 + m_2 = 2m_1$ and

$$2^{m_2}x_0 = x_1(2^{m_1} + 1) - x_2, \quad x_0 = x_1(2^{m_1} + 1) - 2^{m_0}x_2.$$

Thus

$$2^{m_2}x_0 = x_1(2^{m_1} + 1) - x_2 = 2^{m_2}(2^{m_1} + 1)x_1 - 2^{m_2+m_0}x_2,$$

implying that

$$(2^{m_2} - 1)(2^{m_1} + 1)x_1 = (2^{m_0+m_2} - 1)x_2 = (2^{2m_1} - 1)x_2.$$

Hence

$$(2^{m_2} - 1)x_1 = (2^{m_1} - 1)x_2$$

and $s_1s_2 = s_2s_1$, a contradiction.

We conclude that

$$|S^2| \geq |\{s_0^2, s_0s_1, s_0s_2, s_2s_0\} \setminus S'^2| + |S'^2| \geq 3 + 4 = 7 > 3.5|S| - 4.$$

The proof of Lemma 7 is complete. □

Proof of Theorem 7.

Let S be a finite set satisfying the assumptions of Theorem 7. Inequality

$$|S^2| < 3.5k - 4$$

and Lemma 7 imply that

$$S = S_0 = b^{m_0}a^{A_0},$$

where $m_0 \geq 0$.

The set S is non-abelian, so $m_0 \geq 1$. If $m_0 \geq 2$, then Corollary 2 implies that $|S^2| \geq 4k - 4 > 3.5k - 4$, which contradicts our hypothesis. Therefore

$$S = S_0 = ba^A.$$

Theorem 7 now follows from Theorem 5. □

REFERENCES

- [1] B. BUKH, Sums of dilates, *Combin. Probab. Comput.* **17** (2008), no. 5, 627–639.
- [2] J. CILLERUELO, Y. O. HAMIDOUNE, O. SERRA, On sums of dilates, *Combin. Probab. Comput.* **18** (2009), no. 6, 871–880.
- [3] J. CILLERUELO, M. SILVA, C. VINUESA, A sumset problem, *J. Comb. Number Theory* **2** (2010), no. 1, 79–89.
- [4] G. A. FREIMAN, Foundations of a structural theory of set addition, *Translations of mathematical monographs*, **37**, Amer. Math. Soc., (1973), Providence, Rhode Island.
- [5] G. A. FREIMAN, M. HERZOG, P. LONGOBARDI, M. MAJ, Small doubling in ordered groups, *J. Austral. Math. Soc.* (to appear).
- [6] B. GREEN, What is ... an approximate group ?, *Notices Amer. Math. Soc.* **59** (2012), no. 5, 655–656.

- [7] Y. O. HAMIDOUNE, J. RUÉ, A lower bound for the size of a Minkowski sum of dilates, *Combin. Probab. Comput.* **20** (2011), no. 2, 249–256.
- [8] V. F. LEV, P. Y. SMELIANSKY, On addition of two distinct sets of integers, *Acta Arith.* **70** (1995), no. 1, 85–91.
- [9] M. B. NATHANSON, Inverse problems for linear forms over finite sets of integers, *J. Ramanujan Math. Soc.* **23** (2008), no. 2, 151–165.
- [10] D. SHAN-SHAN, C. HUI-QIN, S. ZHI-WEI, On a sumset problem for integers, *arXiv:1011.5438* (2010)
- [11] Y. V. STANCHESCU, On addition of two distinct sets of integers, *Acta Arith.* **75** (1996), no. 2, 191–194.
- [12] Y. V. STANCHESCU, On the structure of sets with small doubling property on the plane (I), *Acta Arith.* **83** (1998), no. 2, 127–141.
- [13] Y. V. STANCHESCU, The structure of d-dimensional sets with small sumset, *J. Number Theory* **130** (2010), no. 2, 289–303.
- [14] T. C. TAO, Product set estimates for noncommutative groups, *Combinatorica* **28** (2008), no. 5, 547–594.

GREGORY A. FREIMAN

SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, TEL AVIV 69978, ISRAEL.

E-mail address: grisha@post.tau.ac.il

MARCEL HERZOG

SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, TEL AVIV 69978, ISRAEL.

E-mail address: herzogm@post.tau.ac.il

P. LONGOBARDI

DIPARTIMENTO DI MATEMATICA E INFORMATICA, UNIVERSITA DI SALERNO, 84084 FISCIANO (SALERNO), ITALY

E-mail address: plongobardi@unisa.it

M. MAJ

DIPARTIMENTO DI MATEMATICA E INFORMATICA, UNIVERSITA DI SALERNO, 84084 FISCIANO (SALERNO), ITALY

E-mail address: mmaj@unisa.it

YONUTZ V. STANCHESCU

THE OPEN UNIVERSITY OF ISRAEL, RAANANA 43107, ISRAEL

and

AFEKA ACADEMIC COLLEGE, TEL AVIV 69107, ISRAEL.

E-mail address: ionut@openu.ac.il and yonis@afeka.ac.il